

Douglas Felipe

 dougbiomed@gmail.com

<https://datadesigner.pro> | **Link:** <https://www.linkedin.com/in/douglas-felipe/>

Resumo

Biomédico e pesquisador com experiência em Bioinformática e Biotecnologia, especializado em Data Science, Inteligência Artificial e Inteligência Computacional pelo Instituto Metrópole Digital. Atualmente, sou graduando em Engenharia de Software e doutorando em Bioinformática, atuando em projetos e pesquisas nos campos da Bioinformática, Ciência de Dados, IA e Machine Learning. Tenho experiência em projetos multidisciplinares que combinam inovação tecnológica e impacto prático, resultando em publicações científicas, patentes e premiações. Atuo na análise de dados, construção e automação de pipelines, desenvolvimento de modelos preditivos e de classificação, criação de dashboards e aplicações web interativas, utilizando tecnologias de IA e Machine Learning com Python e R, aliando conceitos de MLOps e DevOps para otimizar fluxos de trabalho e processos de análise. Além da experiência técnica, possuo habilidades em comunicação e produção científica, incluindo a elaboração de relatórios técnicos, documentação de pipelines e algoritmos, garantindo a acessibilidade e reprodutibilidade para diferentes públicos.

Formação Acadêmica

Doutorado: Bioinformática

Universidade Federal do Rio Grande do Norte

Status - Cursando

Bacharelado: Engenharia de Software

Universidade Federal do Rio Grande do Norte

Status - Cursando

Mestrado: Bioinformática

Universidade Federal do Rio Grande do Norte | 2023

Status - Concluído

Bacharelado: Biomedicina

Universidade Federal do Rio Grande do Norte | 2019

Status - Concluído

Formação Complementar

- **Ciência de Dados (360h)** - Instituto Metr pole Digital (UFRN)
- **Intelig ncia Artificial (360h)** - Instituto Metr pole Digital (UFRN)
- **Intelig ncia Computacional (360h)** - Instituto Metr pole Digital (UFRN)
- **Machine Learning Scientist in Python (86h)** - DataCamp
- **Data Analyst with Python (36h)** - DataCamp
- **Data Analyst with R (36h)** - DataCamp
- **SQL Fundamentals (26h)** - DataCamp
- **Supervised Machine Learning in Python (25h)** - DataCamp
- **Machine Learning Fundamentals with Python (16h)** - DataCamp
- **Data Visualization in R (12h)** - DataCamp
- **AI Business Fundamentals (11h)** - DataCamp
- **AI Fundamentals (10h)** - DataCamp
- **PostgreSQL (8h)** - Alura

Habilidades e Compet ncias

- **Data Science e Machine Learning:** Experi ncia em fundamentos, implementa  es e constru  o de pipelines, integrando boas pr ticas de pr -processamento, modelagem e predic  o.
 - **Linguagens de Programac  o:** Profici ncia em **Python** e **R**, aplicados   ci ncia de dados, aprendizado de m quina e visualiza  o de dados.
 - **Manipula  o e An lise de Dados:** Experi ncia com ferramentas como **Pandas**, **NumPy**, **dplyr** e **tidyr** para limpeza, transforma  o e an lise de dados.
 - **Visualiza  o de Dados:** Habilidade no uso de bibliotecas como **Matplotlib**, **Seaborn**, **Plotly**, **ggplot2**, **Dash** e **Shiny** para criar gr ficos interativos e dashboards.
 - **Estat stica e C lculo Aplicados:** Conhecimento em estat stica descritiva, inferencial e c lculo para suporte a an lises.
 - **Boas Pr ticas de Desenvolvimento:** Experi ncia com **DevOps**, **MLOps**, controle de vers o (**Git**, **GitHub**, **GitLab**) e utiliza  o de containers com **Docker**.
 - **Documenta  o T cnica:** Conhecimento de padr es UML para cria  o de diagramas e desenvolvimento de documenta  o t cnica estruturada.
 - **Comunica  o T cnica:** Capacidade de traduzir informa  es complexas em apresenta  es claras e acess veis, tanto para p blicos t cnicos quanto n o t cnicos.
 - **Cria  o de Relat rios T cnicos:** Gera  o de relat rios utilizando ferramentas como **RMarkdown**, **Quarto** e **LaTeX**.
 - **Desenvolvimento de Solu  es Personalizadas:** Habilidade em implementar solu  es adaptadas  s necessidades espec ficas em projetos envolvendo ci ncia de dados e machine learning.
-

Experiência Profissional

Pesquisador

Laboratório de Biologia Molecular e Genômica | UFRN | Setembro 2020 - atual

- Realizei análises genômicas completas, desde o pré-processamento até a montagem e anotação funcional, conduzindo estudos comparativos de genomas de isolados bacterianos. Utilizei bancos de dados biológicos (KEGG, NCBI, CAZy, PubChem) e técnicas de Big Data para processamento e armazenamento eficiente de dados, empregando plataformas de análise genômica como KBase e Galaxy, além de pipelines automatizados em Python e R.
- Desenvolvi uma aplicação web interativa para processamento e visualização de dados genômicos, utilizando Python e o framework Dash. A aplicação foi implantada na plataforma Digital Ocean, garantindo acessibilidade a pesquisadores interessados na análise de genomas com foco no potencial de biorremediação de bactérias, fungos e plantas.

Pesquisador

Laboratório de Investigação do Câncer e Inflamação | UFRN | Março 2024 - Setembro 2024

- Atuei na análise computacional de biomarcadores tumorais utilizando abordagens de Deep Learning (SignalP, NetMHCpan, NetMHCpan II) para identificar alvos específicos no câncer colorretal. Desenvolvi e implementei pipelines de Data Science em Python e R, otimizando a seleção de peptídeos imunogênicos com potencial para aplicação em nanovacinas.

Pesquisador

Microciclo Biotecnologia LTDA | Janeiro 2024 - Agosto 2024

- Atuei na caracterização de isolados bacterianos para o desenvolvimento de consórcios biorremediadores aplicados à descontaminação de resíduos oleosos industriais. Utilizando Python e R, realizei análises genômicas que auxiliaram na caracterização genética dos isolados empregados na formulação de uma solução biotecnológica patenteada (BR 10 2021 0242540), capaz de remover até 78% dos contaminantes em apenas 7 dias. Essa inovação oferece uma alternativa sustentável para empresas, reduzindo custos operacionais e impactos ambientais no descarte de resíduos contaminados.

Pesquisador

Laboratório de Biologia Molecular e Genômica | UFRN | Agosto 2017 - Julho 2019

- Conduzi estudos exploratórios e preditivos com dados do TCGA e KEGG para investigar a associação entre tabagismo e subtipos moleculares do câncer de bexiga. Apliquei pipelines de Data Science e Bioinformática para processamento, integração e análise dos dados, identificando biomarcadores genéticos correlacionados à redução da sobrevida dos pacientes em até 60 meses.

Idiomas

Português: Língua materna

Inglês:



Proficiente

Espanhol:



Intermediário